



PCT09

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/980,559

DATE: 02/06/2003

TIME: 13:51:03

Input Set : A:\20264301.app

Output Set : N:\CRF4\02052003\I980559.raw

```

3 <110> APPLICANT: Yanaqi, Masayuki
4      Emerson, Suzanne
5      Pukh, Jens
6      Purcell, Robert
8 <120> TITLE OF INVENTION: Cloned Genome of Infectious Hepatitis C Viruses of
9      Genotype 2a and Uses Thereof
11 <130> FILE REFERENCE: 20264302PC
C--> 13 <140> CURRENT APPLICATION NUMBER: US/09/980,559
14 <141> CURRENT FILING DATE: 2010-06-02
16 <150> PRIOR APPLICATION NUMBER: 60/137,693
17 <151> PRIOR FILING DATE: 1999-06-04
19 <160> NUMBER OF SEQ ID NOS: 39
21 <170> SOFTWARE: PatentIn Ver. 2.1
23 <210> SEQ ID NO: 1
24 <211> LENGTH: 3711
25 <212> TYPE: DNA
26 <213> ORGANISM: Hepatitis C virus
28 <400> SEQUENCE: 1
29 aacccgccct aatagggggg acactcggcc atgaatract cccctgtgag gaactactgt 60
30 attcaagcag aaagcgrcta gcratggcgt taglatgagt gtcgtacagc ctccaggccc 120
31 cccctctccg ggagagccat agtggtctgc ggaacccgtg agtacacagg aattgccggg 180
32 aagaactggg catttcttg ataaaccac tatatgccg gccatttggg cgtgcccncg 240
33 aaagactgct agccagtag cgttgggttg cgaaggccct tctggtactg cctgataggg 300
34 tgcctggcag tgcncggga ggtctcgtag accgtgcacc atgagcacia atccataaac 360
35 tcaaaagaaa accaaaagaa acacaaaccg tgcacacaa gacgttaagt tcccgggcgg 420
36 cggccagata gtggcgagg tatacttggt ccgcgcagg gcccccagg tgggtgtgcg 480
37 ccgacaaagg aagacttcgg agrggtcna gcacgttga aggcgcagg ccatacccaa 540
38 agatcggcgc tccactggca aatcttggg aaaaaccagg tacccttggc cctatabcgg 600
39 gaatgaggga ctggcgagg caggatggct cctgtccccc cgaggttccc gtccctcttg 660
40 gggccccaat gaccccggc ataggctcgg caacgtgggt aaggctcctg ataccctaac 720
41 gtgcgccttc gcgacctca tggggtagat cctgttcgtg ggcccccgc tggcgggcgt 780
42 cgcacagagt ctgcgcctg gctgagagt cctggaggac ggggttaatt ttgcacagg 840
43 gaacttaacc ggttgcctt ttctatctt ctgcctggc ctgctgtct ccataccac 900
44 ccgggtctcc gctgcgagg tgaagaacat cagtaaccgc tacatggtga ctacagactg 960
45 taccaatgac agcattacct ggcagctcaa ggtgtgtg ctccacgtcc ccgggtggt 1020
46 ccgtgaggag aaagtgggga atgcattctt gtgttgata ccggtctaac cgaatgtggc 1080
47 cgtgcagcgg ccggcgccc tcaagcaggg cttggcgacg cacatcgaca tggttgtgat 1140
48 gtccgcacag ctctgtctg cctctacgt gggggacctc tgggtggg tcatgtctg 1200
49 agcccaaatg ttcattgtct cggcgacgca ccactgatt gtccaagact gcaattgtc 1260
50 catctacctt ggtaccatca ctggacaccg catggcatgg gacatgatga tgaactgttc 1320
51 gccacagatt accatratct tgggtacgc gctcgtgtc ccgaggttca ttatagacat 1380
52 cattagcggg gctcatggg gctcatgtt cggctggcc taattctct tgcagggagc 1440
53 gtaggagaaa gtgtgtgtc taattctgtt gggcgagg gtggagggc gcaacacata 1500

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/980,559

DATE: 02/06/2003

TIME: 13:51:0

Input Set : A:\20264301.app

Output Set: N:\CRF4\02052003\I980559.raw

```

14 tgttgggggt tctccggcgc aqacacagg ggcctccacc apcttattttg acatgggcgc 1341
15 caggcagaaa atccagctcg ttaacacaa tccagctcg ccacacacc gacacgcct 1342
16 ggaatgaat gactcttcg aacacacat tatcgctct ctattctaca cccacacatc 1343
17 cactcgcca gactgtccg aacacatct tgcctgccc agtatcagg ccttccgggt 1344
18 ggaatgggc gacttgaat atpccatca tcttctaat ccagaggata tgaacacct 1345
19 ttctggcac taccacccaa ggcctctcg cgtctctct ccacagactg tgtgtgggc 1346
20 agctgactct ttccacccaa ggcctctct agtggcagg ccagacaggc ttccaggggc 1347
21 cacttccagc tggggggaga atpccatca tcttctaat ttccacagca ctccacacac 1348
22 gctgggctca tggctgggt ggcctctct gacctctct agtccacca apacttggg 1349
23 caccacccc tggctgaat pccatctct ccagcagc ccagacctgt tgtgcccac 1350
24 gacctctct aggcagcct ctccacacac ctacacaca tggcgtctcg cgcctgggc 1351
25 caggccagg tgcctgctcg actacacct ccagctctcg cactacacct gacacgtta 1352
26 ctacacacat ttccacacaa ggcctctct ggcagggtt ccacacaggc tccaggggt 1353
27 atpccatct cactgtggg atgttgcac ctctggagg ccagacagaa gtcactggc 1354
28 cactgtgctt actccacca ggcctctct cacttctct gctctctct ccgactggc 1355
29 cgcctgtct actgtctct cccacacaa cccacacat cccacacat cactcagta 1356
30 tggcctctca cctgcctca cccacacat cgcctgggc agtggttaa cactcagta 1357
31 cctgctctta ggcacagca ggcctctct cgtcttgg atgtctctct tcttgggc 1358
32 ggcagaggc gactcagca ggcctctct ctccacacat ccagcagc ctccagcag 1359
33 tggcctctca tcttctctca tcttctctct cgcctgggc ccagcagc ctccagcag 1360
34 cgcctctct actacacac cccacacat ggcctctct ccagcagc ctccagcag 1361
35 ggcacacag ggcctctct cccacacat cgcctctct ccagcagc ctccagcag 1362
36 ggcacacag ggcctctct cccacacat cgcctctct ccagcagc ctccagcag 1363
37 ggcacacag ggcctctct cccacacat cgcctctct ccagcagc ctccagcag 1364
38 ggcacacag ggcctctct cccacacat cgcctctct ccagcagc ctccagcag 1365
39 ggcacacag ggcctctct cccacacat cgcctctct ccagcagc ctccagcag 1366
40 ggcacacag ggcctctct cccacacat cgcctctct ccagcagc ctccagcag 1367
41 ggcacacag ggcctctct cccacacat cgcctctct ccagcagc ctccagcag 1368
42 ggcacacag ggcctctct cccacacat cgcctctct ccagcagc ctccagcag 1369
43 ggcacacag ggcctctct cccacacat cgcctctct ccagcagc ctccagcag 1370
44 ggcacacag ggcctctct cccacacat cgcctctct ccagcagc ctccagcag 1371
45 ggcacacag ggcctctct cccacacat cgcctctct ccagcagc ctccagcag 1372
46 ggcacacag ggcctctct cccacacat cgcctctct ccagcagc ctccagcag 1373
47 ggcacacag ggcctctct cccacacat cgcctctct ccagcagc ctccagcag 1374
48 ggcacacag ggcctctct cccacacat cgcctctct ccagcagc ctccagcag 1375
49 ggcacacag ggcctctct cccacacat cgcctctct ccagcagc ctccagcag 1376
50 ggcacacag ggcctctct cccacacat cgcctctct ccagcagc ctccagcag 1377
51 ggcacacag ggcctctct cccacacat cgcctctct ccagcagc ctccagcag 1378
52 ggcacacag ggcctctct cccacacat cgcctctct ccagcagc ctccagcag 1379

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/980,559

DATE: 02/06/2003

TIME: 13:51:03

Input Set : A:\20264301.app

Output Set : N:\CRF4\02052003\I980559.raw

```

103 ggaagtgagg cttggggcagg agggcgagat ccccttctat gggaggggga ttccctgttc 4500
104 ttacutcaag ggaaggaagc atc gatctt ctgccattca aagaaaaggt dtgacqader 4560
105 cgcggggggc cttgggggta tgggcttgaa ct cagtggca tactacagag ggttggaggt 4620
106 ctccgttaaa ccaactcagg gagaagtatg ggtcttgcgc accgaagccc tcatgacaag 4680
107 gtatactggg gactttgact ccttgatcga ctgcaacgta ggggtcactc aagttgtaga 4740
108 cttcagttta gacccacat tcaccataac cacacagatt gtcctcgaag acgtgtcttc 4800
109 aegt agcag cgcgggggtc gcaagggt ag ggaagagctg ggcatttata ggtatgttcc 4860
110 cactggtgag cgaagcctcag gaatgtttga cagtgtagtg ctctgtgagt gctacgaagc 4920
111 agggggcgca tggatagagc tcacacacac ggagaccacc gt caggetca gggcgtatct 4980
112 caaacagccc ggtttgcttg tgtgcaaga ccatcttgag ttctgggagg cagttttcgc 5040
113 cggcctcaca cacatagatg cccacttctt ttcccaaca aagcaatcgg gggaaaaatt 5100
114 cgcataacta acagcctacc aggtacaggt gtgggttagg gccaaagccc cccccccgtc 5160
115 ctgggaagtc atgtggaggt gtttgactcg actcaagccc aaactgttgg gcccacacac 5220
116 tctcctgtac cgtttgggt ctgttaacca cgaaggtcac ctacacacac cctggaagca 5280
117 ataatcgcc cctgtgatgc aagcggacct tgaggctcat accagcaaat gggctttgjc 5340
118 agggggagtc ctggggggcg tggcgggtta ttgctggcg accgggtgtg ttgtcatcat 5400
119 cggcggcttg cacattaacc agcagagcgt cgttggcgcg gacaaggagj tctctatja 5460
120 ggtttttgat gagatggag aatgtgcctc taggggggt ctcatgaaq aggggcagcg 5520
121 gatagcgag atgtgaggt ccaagatcca aggttattg cagcaagctt ccaaacaaag 5580
122 tcaagacata caaccacact tgcagggttc atgggccaag gttagaacaat tctgggcccac 5640
123 acacatgtgc aacttcttta ggggaatcca atactcgca ggaattaaa cactgcaagj 5700
124 gaacccctga gtactttcca tgatgggtca cagtggcgcc ctacacagtc cgtgtcaac 5760
125 aagcaccact atccttctca acattttggg ggggtggcta gcattccaaa ttgcacacac 5820
126 cggggggggc attggtctcg ttgtcaatgg cctagtggga gctgongtag gtagtatagg 5880
127 cttaggttag gctgtagtgg acatcctggc aggttatggt gcggggcatt cgggggtct 5940
128 cgtgcatttc aagatcatgt ctggcgagaa gcccctccat gaggatgtcg tcaacttqct 6000
129 gcttggaatt ctgtctcagg gtgccttgggt agtggagtc atctggggg cactctcgcg 6060
130 cggacacgtg ggaacggggg aaggcgaggt ccaatggatj aatagactca ttgpccttjc 6120
131 tccagaggga aatcagctcg ccccaaccca ctacgtgacg gagtgggatj cgtcgpagcg 6180
132 cgtgacccaa ctactgggt cccctacacat aaccagcgtj ctacgaagj tccacacact 6240
133 gattactgag gactgcccc tcccatcgcg cggctcgttg ctccgggatj tgtgggaact 6300
134 ggtttgcacc atcctaacag actttaaaaa ttgggtgacc tccaaattat tccaaagat 6360
135 gcccggtctc cctttgtct cctgtcaaaa ggggtacaag ggcgtgtggj cgggactgg 6420
136 catcatgacc acacgggtgc cttgggggcg caatatctct ggcaatgtcc gttggggtct 6480
137 catgagaatc accggggcta agacctgcac gaatatctgg caggggaact ttcctatcaa 6540
138 ttgttacacg gagggcaggt gcttgccgaa accgggcaca aactttaagg tggccatcg 6600
139 gagggtggcg gctcagagt acgaggaggt gacgpagac gggctatacc actacataac 6660
140 aggaactcac actgataact tgaaggtccc ctgcaacta cccctcccg agttcttttc 6720
141 ctgggtggac gtagtgagga tccataggtt tgcccccaca ccgaagcgtt ttttcgggja 6780
142 cgaggtctcg ttctggttg ggttaatttc attgtgtctc gggctccagc ttctttgga 6840
143 ccttgaaacc gacacagagc tattgatgtc catgttaaca gatccatctc atatcagggc 6900
144 ggagactgca ggcggggttt tagcgcgggg gtaccccca tccgaggcaa gctcctcgjc 6960
145 gagccagjta tggcaccat gcttgagagc cactgcacc acccagggca aagcctatja 7020
146 tgtggacatg gtgagctcta acctgtctat ggggggggat gtgactcgga tagagctcg 7080
147 gtccaaagtg gtcgtctcg acctctctga ccaatggtc gaagaaagja ggcacttja 7140
148 gcttctgata ccatcagaat acatgtctcc caagaagagg ttcccaacag ctttacccgc 7200
149 ctgggcacgg cctgattaca accacagat tctgggaatg tggaaaaggc cagattacca 7260
150 accggccact gttgggggt gtgtctctcc tctctatagg aaaaacccga cgcctcccc 7320
151 aagqagggcg ccgacagtgc gctaaagtga ggactcacta ggagatgcac ttcaacagct 7380

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/980,559

DATE: 01/06/1998

TIME: 13:51:03

Input Set : A:\20264301.app

Output Set : N:\CRF4\02052003\I980559.raw

```

152 ggcattaaag tcttttggcc agcccccrcr aagcggcgat tcaggccctt ccacqggggc 7440
153 qqqeqctgcc gattccggca gtcagacgc tctgatgag ttggcccttt cggagacagg 7500
154 ttcctctctt tccatgcccc ccttcgaggg ggagcttggg gatccagacc tggagcctga 7560
155 qdaggtagag ccccaacccc ccccccaggg ggggggtggc gctcccggtt cggactcggg 7620
156 qtcctggtct acttgcctcg aggaggaagg ctccgtcgig tctgctcca tgcataactc 7680
157 ctggacccgg gctctaataa ctccctttag tccogaagag gagaagttac cgtattaacc 7740
158 cttagcaaac tccctgttgc qatatcacaa caagggttac tgaaccacaa caaagagcgc 7800
159 ctcaaaagg gctaaaaagg taactttga taagatgcaa gtgctcgact cctactacga 7860
160 ctcaactcta aaggacatta agctagcggc ctccaaggte accgcaaggc tctcaccat 7920
161 ggaqjaggtt tgcagtttaa cccaccccc ttctgcaaga tctaaatatg ggtttggggc 7980
162 aaqjaggtc cgaagcttgt cggggagggc cgttaaccac atcaagtcct tgtggaagga 8040
163 cctcctggag gactcagaaa caccatttc cacaaccact atggcctaaa atgaggtgtt 8100
164 ctgcctggac cccaccaagg ggggcaagaa agcagctcgc ctatcgctt accctgacct 8160
165 cggcctcagg gctcgggaga agatggccct ttatgacatt acacaaaaac ttctcagggc 8220
166 ggtgatgggg gcttctatag gattccagta ttcccccgtt cagcgggtag agtttctctt 8280
167 gaaaactatg ggggaaaaga aggaacctat ggggttttgg tatgataccc gatgctttga 8340
168 ctcaaccgtc actgagagag acatcaggac tgaggagtcc atatatcggg cctgctcctt 8400
169 gcccjaggaq gcccacactg ccatacactc gctaacctag agactttacg tgggagggcc 8460
170 tatgttcac agcaagggcc aaacctggcg gtacaggcgt tgcggggcca gcggggtgt 8520
171 aacctatgg atgggjaaca ccatacactg ctactgaaa gccttagcg ctgttaaagc 8580
172 tgcagggata atgcggccca caatgctgtt atggggcgat gacttggttg tctctcaga 8640
173 aagcagggg acgagjagg acgaggggaa cctgagagcc ttcaggagjg ctatgaccag 8700
174 gtatctctgc cctccgtgtg acccccbaq accggagtat gatctggagc tgataacatc 8760
175 atgctctcca aatgtgtctg tggggtggg cccaccaagg cgcggcagat actacctgac 8820
176 cagaiaacct accactccaa tggccgggg tgcctgggaa acagttcgac actccctgt 8880
177 caattcatgg cggggaacaa tctatcagta cgcctcgacc atatgggtc gcatggctct 8940
178 gatgacacac tctttctcca ttctatggc caagacacg ctggaccaga acctcaactt 9000
179 tgagatgtac ggagcgtgtt actccgtgag tcccttggac ctcccagcta taattgaaag 9060
180 gtaacatggg cttagcgtt ttctctcgca cacatacact ccccacgaa tgacacgggt 9120
181 ggttcagacc ctacagaaaac ttggggggcc accctcaga gctggaaga gccgggacg 9180
182 tgcagtcagg ggttcctcca tctccgtgg ggggagagcg ggcgtttgcg gtcgatatct 9240
183 ctcaatttgg ggggtgaaga ccaagctcaa actcactcca ttgcgggaag cggcctctct 9300
184 ggaattatcc agctggttca cctcggggc cggcggggg jacatttatc acagcgtgtc 9360
185 ggtgcccga ccccgcttat tgcctttgg cctactcta cttttttag ggttaggctt 9420
186 ttctctact cccgtctggt agagcggcac acattagcta cactccatag ctaactgtcc 9480
187 ttttttttt ttttttttt ttttttttt ttttttttt ctttttttt ttttttttt 9540
188 ttttttttt ttttttttt ttttttttt ttttttttt ctttttttt ttttttttt 9600
189 ttttttttt cctggtggct ccattttag cctagtcacg gctagctgtg aaaggtccgt 9660
190 ggcgcgatg actgcagaga gtgcgtaac tggctctctt gcagatcatg 9711

```

191 210 SEQ ID NO: 2

194 211 LENGTH: 3033

197 212 TYPE: PRT

198 213 ORGANISM: Hepatitis C virus

199 400 SEQUENCE: 2

```

199 Met Ser Thr Asn Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn
201 1 5 10 15
202 Arg Arg Pro Gln Asp Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly
203 20 25 30
205 Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Ala

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/980,559

DATE: 02/06/2003

TIME: 13:51:03

Input Set : A:\20264301.app

Output Set: N:\CRF4\02052003\I980559.raw

```

106          35          40          45
111 Thr Arg Lys Thr Ser Glu Arg Ser Gln Pro Arg Gly Arg Arg Gln Pro
114          50          55          60
117 Ile Pro Lys Asp Arg Arg Ser Thr Gly Lys Ser Thr Gly Lys Pro Gly
120          65          70          75          80
124 Tyr Pro Trp Pro Leu Tyr Gly Asn Glu Gly Leu Gly Trp Ala Gly Trp
127          85          90          95
131 Leu Leu Ser Pro Arg Gly Ser Arg Pro Ser Trp Gly Pro Asn Asp Pro
134          100          105          110
137 Arg His Arg Ser Arg Asn Val Gly Lys Val Ile Asp Thr Leu Thr Cys
140          115          120          125
143 Gly Pro Ala Asp Leu Met Gly Tyr Ile Pro Val Val Gly Ala Pro Leu
146          130          135          140
149 Gly Gly Val Ala Arg Ala Leu Ala His Gly Val Arg Val Leu Gln Asp
152          145          150          155          160
155 Gly Val Asn His Ala Thr Gly Asn Leu Pro Gly Cys Ser Pro Ser Ile
158          165          170          175
161 Phe Leu Leu Ala Leu Leu Ser Cys Ile Thr Thr Pro Val Ser Ala Ala
164          180          185          190
167 Glu Val Lys Asn Ile Ser Thr Gly Tyr Met Val Thr Asn Arg Cys Thr
170          195          200          205
173 Asn Asp Ser Ile Thr Trp Gln Leu Gln Ala Ala Val Leu His Val Pro
176          210          215          220
179 Gly Cys Val Pro Cys Glu Lys Val Gly Asn Ala Ser Gln Cys Trp Ile
182          225          230          235
185 Pro Val Ser Pro Asn Val Ala Val Gln Arg Pro Gly Ala Leu Thr Gln
188          240          245          250          255
191 Gly Leu Arg Thr His Ile Asp Met Val Val Met Ser Ala Thr Leu Cys
194          260          265          270
197 Phe Ala Leu Tyr Val Gly Asp Leu Cys Gly Gly Val Met Leu Ala Ala
200          275          280          285
203 Gln Met Phe Ile Val Ser Pro Gln His His Trp Phe Val Gln Asp Cys
206          290          295          300
209 Asn Cys Ser Ile Tyr Pro Gly Thr Ile Thr Gly His Arg Met Ala Trp
212          305          310          315          320
215 Asp Met Met Met Asn Trp Ser Pro Thr Ala Thr Met Ile Leu Ala Tyr
218          325          330          335
221 Ala Met Arg Val Pro Glu Val Ile Ile Asp Ile Ile Ser Gly Ala His
224          340          345          350
227 Trp Gly Val Met Phe Gly Leu Ala Tyr Phe Ser Met Gln Gly Ala Trp
230          355          360          365
233 Ala Lys Val Val Val Ile Leu Leu Ala Ala Gly Val Asp Ala Arg
236          370          375          380
239 Thr His Thr Val Gly Gly Ser Ala Ala Gln Thr Thr Gly Arg Leu Thr
242          385          390          395          400
245 Ser Leu Phe Arg Met Gly Pro Arg Gln Lys Ile Gln Leu Val Asn Thr
248          405          410          415
251 Asn Gly Ser Trp His Ile Asn Arg Thr Ala Leu Asn Cys Asn Asp Ser
254          420          425          430

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/980,559

DATE: 02/06/2003

TIME: 13:51:04

Input Set : A:\20264301.app

Output Set: N:\CRF4\02052003\I980559.raw

L:13 M:270 C: Current Application Number differs, Replaced Current Application Number